



Lee, T. V. N., Ireland, A. J., Atack, N. E., Deacon, S. A., Jones, T. E. M., Matharu, J., Wills, A., Al-Ghatam, R., Richard, B. M., Ness, A. R., & Sandy, J. R. (2019). Is There a Correlation Between Nasolabial Appearance and Dentoalveolar Relationships in Patients With Repaired Unilateral Cleft Lip and Palate? *Cleft Palate-Craniofacial Journal*. <https://doi.org/10.1177/1055665619862145>

Peer reviewed version

Link to published version (if available):
[10.1177/1055665619862145](https://doi.org/10.1177/1055665619862145)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Sage at <https://journals.sagepub.com/doi/full/10.1177/1055665619862145>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Is there a correlation between nasolabial appearance and dento-alveolar relationships in patients with repaired unilateral cleft lip and palate?

AbstractKey Words

Cleft, 5-Year Olds' Index, Outcome measures, Nasolabial aesthetics.

Objective

To determine if a relationship exists between the aesthetic scores given to photographic records of the nasolabial region of repaired unilateral cleft lip and palate (UCLP) patients and the 5-Year Olds' Index scores of study models for the same subjects.

Design

Retrospective study.

Setting

University of Bristol Dental Hospital, United Kingdom.

Participants

Non-syndromic UCLP patients previously enrolled in the Cleft Care UK (CCUK) Study.

Methods

The CCUK subjects, who had both study models and photographs (frontal and worm's eye view), were identified and their records retrieved. These were rated by two consultants and two senior registrars in orthodontics. The 5-Year Olds' Index was used to score the study models, and at a separate sitting, a 5-point Likert scale was used to score the cropped frontal and worm's eye view photographs of the same children. The results were analysed using intra-class correlation coefficients and Cohen's Kappa.

Main outcome measures

Correlation between the aesthetic scores of the photographic views and the concordant 5-Year Olds' Index scores of the study models.

Results

The intra-class correlation coefficient scores showed very poor agreement between the photographic views and their concordant study models. The level of inter- and intra-rater reliability was strongest when scoring the study models.

Conclusions

There was no agreement between the scores given to various photographic views and their corresponding study models. Scoring the study models using the 5-Year Olds' Index was the most reliable outcome measure for this age group.

Introduction

Cleft lip, with or without palate, (CL±P) is the most common craniofacial anomaly and occurs in approximately in 1 per 1000 live births in Caucasian populations (Goodacre, 2008). Apart from having an impact on appearance, the cleft may affect speech, hearing, and may be associated with psychosocial problems (Mossey et al., 2009). Those born with orofacial clefting need care by a multi-disciplinary team from childhood through to adulthood.

In order to determine whether or not treatment is effective, it is important to have good outcome measures. These are used not only to determine treatment efficacy, but also to ensure continued improvement in standards of care. A diverse range of outcome tools are available in cleft care and this is probably due to the different types of cleft being assessed, the large number of specialties involved in cleft care and the fact that new outcome measures are continually being developed (Jones et al., 2014). Study models are normally taken at 5 years of age as a proxy to assess the effects of primary surgery on growth of the maxilla using the 5-Year Olds' index (Atack *et al.*, 1997a). This index is a subjective assessment, which uses study models to categorise dental arch relationships in relation to a set of reference models which are grouped into five categories. The study models are assessed against the reference models in their vertical, transverse and anterior posterior relationships. A category 1 rating equates to excellent dental base relationship whereas category 5 is a very poor dental base relationship.

Improved facial appearance and the restoration of normal soft tissue function are the principal aims of primary lip repair and so the appearance of the repaired region has a bearing on whether treatment is deemed successful (Asher-McDade et al., 1991). Despite the importance of the appearance of the nasolabial region, it is particularly difficult to assess in a valid and reliable way (Roberts et al., 1991). It can be assessed using one of three views:

- Frontal View - this allows assessment of the nasal and lip form, as well as symmetry and scarring. It is the most commonly used view (Tobiasen, 1987, Shaw, 1981).
- Profile View - this allows assessment from a lateral view and the impact of treatment on the anterior posterior growth can be observed.

- Worm's Eye View – this “up and under” view allows for assessment of the symmetry of the nostrils.

2D photographic imaging, either black and white photographs (Tedesco et al., 1983), colour slides or colour prints (Asher-McDade et al., 1991, Eliason et al., 1991) have been used to assess facial aesthetics. The disadvantage of using photographs, slides or prints is that they present the three-dimensional face in just two dimensions. In addition, any change in lighting, head orientation and the distance of the camera to the subject will all have an effect on the 2D photographs (Mosmuller et al., 2015). As a result standardised photographs are encouraged (Vegter and Hage, 2000). The validity of using photographs as a replacement for live subjects was investigated by Howells and Shaw (1985), who found a moderately high correlation between appearance of photographs and those of living subjects. The advantage of 2D imaging is that it has been in clinical use for a number of years so that outcomes can be effectively measured over time. Shaw *et al.*, (1985) found that features such as the hair and eyes are more influential than the actual anomaly when assessing facial aesthetics, and as a result cropped photographs are often used (Asher-McDade *et al.* 1991). Using such photographs, a number of indices have been developed for the assessment of nasiolabial outcomes, not all of which have been validated. Examples of such Indices include the Asher-McDade System (Asher-McDade et al., 1991), the VLS classification [V-vermillion, L-lip, and S-scar] (Assuncao, 1992), the Aesthetic Index (Johnson and Sandy, 2003) and the 5 point Likert scale (Okkerse et al., 2001). These indices have their advantages and disadvantages. The 5 point Likert scale (Okkerse et al., 2001) was chosen for this study, this is a 5 point ordinal scale where a score of 1 is excellent and a score of 5 poor.

Following the recommendations of the CSAG study (Williams et al. 2001) there are now 11 centralised cleft units in the UK which provide cleft care via a Multi-Disciplinary Team model. In 2011 the National Institute for Health Research funded the CCUK study where the principal aim was to evaluate the effects of the centralisation of cleft care following CSAG. The subjects included in the study were 5-year-old children born with non-syndromic complete UCLP. It was largely a repeat of the CSAG study and the study models and photographic records from the CCUK Study were used in the current study.

Taking impressions of 5-year olds can be a challenge (Clark *et al.*, 2007), and may be easier to obtain photographic records. The aim of this study was therefore to assess whether the scores given to the nasolabial (frontal and worm's eye) appearance of non-syndromic complete UCLP cleft children correlated with their 5-Year Olds' Index scores for the study models of the same children (Atack *et al.*, 1997b; Atack *et al.*, 1997a). If the results suggested that these correlated then this might encourage a change in clinical practice towards the use of photographs rather than study models to assess the effects of primary surgery.

Method

All of the photographic and study model records from the CCUK study were made available for this study. 250 subjects had complete photographic records and 198 had study model records. 181 had both study models and the corresponding frontal and worm's eye view colour photograph, and it was the records of these children that were retrieved for use in the current study. The photographs had been previously cropped and standardised (Al-Ghatam *et al.*, 2015) and were copied onto an encrypted NHS USB stick for use in the current study. Ethical approval (REC reference number: 10/H0107/33 South West 5 REC) and local research and development approvals had been previously obtained as part of the CCUK study (Persson *et al.*, 2015)

The function RANDBETWEEN was used in Microsoft Excel (Redmond, Washington, USA) to generate random ID numbers to anonymise the original CCUK ID numbers for use in this study. This was done to reduce bias towards different cleft centres by the raters when scoring.

The identified study models were assessed using the 5-Year Olds' Index (Atack *et al.*, 1997b, Atack *et al.*, 1997a) and were scored in two rounds over a five week period, with a two week gap between round one and round two. All 181 study models were scored in round 1, and 50% (90 study models) re-scored in round two. The CCUK ID number was blocked out during scoring and the models were only identifiable by the randomised ID number. This ID was kept the same for both scoring sessions.

The raters used in this study included:

- 2 orthodontic consultants experienced in cleft care and already calibrated to use the 5-year olds index
- 2 senior registrars in orthodontics. They were calibrated to use the 5-Year Olds' Index (Atack et al., 1997b, Atack et al., 1997a) as part of this study.

In order to assess the frontal and worm's eye view colour photographs, they were uploaded onto the Birmingham Institute of Paediatric Plastic Surgery (BIPPS) website, keeping the two views separate (Figure I). The raters were then each given a personalised token (password), which enabled them to access the BIPPS website. A different token (password) was emailed to the raters before round two commenced. The photographs were scored using a 5-point Likert scale (Table I), and all frontal view photographs had to be scored before raters were able to score the worm's eye view photographs. Once a photograph was scored the rater was unable to return and change their score. The raters did have the opportunity to save their unfinished survey and resume at a later time. The first round was open for three weeks, during which time the raters scored 181 frontal view photographs and 181 worm's eye view photographs of the same subjects. After a two week break the second round opened. In the second round 90 frontal views and 90 worm's eye views of the original photographic set were re-scored to assess inter- and intra-rater reliability. The order of the photographs was again randomised for round two, keeping the frontal and worm's eye views separate.

Scoring of the photographs took place at a separate time to the scoring of the study models. This was more convenient for the raters and their job commitments.

All data for both the study model and photograph ratings were inputted into a Microsoft Excel Spreadsheet (Redmond, Washington, USA, Version 2013). The data was checked twice for consistency and corrections made as necessary. The data was analysed using Stata version 14 (StataCorp, College Station, Texas, USA) and is presented in terms of summary agreement scores and intra-class correlation coefficients, with the level of agreement based on that described by (Cicchetti, 1994) . Inter-rater agreement was also determined using


Cohen's Kappa and interpreted using the accepted levels of agreement described by Landis and Koch (1977).

BIPPS
BIRMINGHAM INSTITUTE OF PAEDIATRIC PLASTIC SURGERY

POST UCLP REPAIR SCORING SURVEY

0% 100%

Please rate the following image by using the options below.



1 = Excellent
2 = Good
3 = Fair
4 = Poor
5 = Very Poor

Choose one of the following answers:

- ☐ 1 = Excellent Aesthetic Outcome
- ☐ 2 = Good
- ☐ 3 = Fair
- ☐ 4 = Poor
- ☐ 5 = Very Poor

Figure I- Uploaded cropped image for scoring on the BIPPS Web based platform.
(Reproduced with permission from the Birmingham Institute of Paediatric Plastic Surgery).

Results

The records of 181 matched pairs of 5-year olds' study models and photographs from CCUK were analysed. The photographs comprised both frontal views and worm's eye views and the sample comprised the matched records of 125 Males and 56 Females.

The intra-examiner agreement for each of the three sets of records, frontal view, worm's eye views and study models are illustrated in Table I.

Table I- The intra-class correlation and 95% confidence intervals for each of the raters for intra-examiner agreement using the frontal view photographs, worm's eye view photographs and study models.

Rater ID	ICC (95% Confidence Interval)		
	Frontal	Worm's eye view	Study Models
Rater 1	0.71 (0.58 to 0.80)	0.59 (0.33 to 0.74)	0.93 (0.90 to 0.96)
Rater 2	0.66 (0.52 to 0.76)	0.60 (0.36 to 0.75)	0.88 (0.83 to 0.92)
Rater 3	0.64 (0.50 to 0.75)	0.75 (0.57 to 0.85)	0.94 (0.92 to 0.96)
Rater 4	0.53 (0.36 to 0.67)	0.57 (0.41 to 0.59)	0.93 (0.90 to 0.95)

For the frontal views three out of the four raters showed good intra-examiner agreement when assessing the frontal view photographs over the two-time periods. Rater 4 showed a fair level over the two-time periods. For the worm's eye view one rater (rater 3) showed excellent agreement, one rater had good agreement and the remaining two had fair agreement over the two-time periods.

When scoring the study models using the 5-Year Old's Index (Atack et al., 1997b, Atack et al., 1997a) all four raters showed excellent agreement over the two-time periods.

The intra-class correlation coefficients were determined for agreement between each of the photographic types and the scores for the study models (see Table II).

Table II- Comparison of the intra-class correlation and 95% confidence intervals for agreement between scores for the frontal and study models versus worm's eye view and study models.

	ICC Value - Frontal view vs Study Models - (95% CI)	ICC Value- Worm's eye view vs Study Models – (95% CI)
Rater 1	0.03 (-0.11 to 0.18)	0.04 (-0.10 to 0.19)
Rater 2	0.02 (-0.13 to 0.16)	0.02 (-0.13 to 0.16)
Rater 3	0.04 (-0.11 to 0.18)	0.03 (-0.12 to 0.17)
Rater 4	0.10 (-0.04 to 0.25)	0.12 (-0.03 to 0.26)

There was very poor agreement between the scores for both the frontal photographs and study models, and similarly between the worm's eye view photograph scores and study model scores.

Comparing whether worm's eye view photographs had better agreement to study models, than frontal views to study models the results varied between raters (Table II). Two raters (1 and 4) showed better agreement with worm's eye view, rater 3 showed worse agreement with worm's eye views and rater 2 had the same level of agreement irrespective of the view.

Inter-rater reliability for the two Consultant Orthodontists experienced in the use of the 5-Year Olds' Index (Atack et al., 1997b, Atack et al., 1997a) was assessed using both the intra-class correlation coefficient and also Cohen's Kappa. Using Cohen's Kappa with the frontal photographic views there was fair agreement, for the worm's eye view there was moderate agreement and for the study model scores substantial agreement (Table III) and this is mirrored for the intra-class correlation coefficient.

Table III – Inter-examiner agreement for frontal view, worm's eye view and study models

	ICC (95% Confidence Interval)	Kappa
Frontal view	0.50 (0.39 to 0.60)	0.28
Worm's eye view	0.68 (0.59 to 0.75)	0.50
Study models	0.87 (0.83 to 0.90)	0.74

Inter-rater reliability for the two registrars was assessed using both the intra-class correlation coefficient and also Cohen's Kappa as shown in Table IV. Using the Cohen's Kappa, the frontal photographic views had fair agreement, the worm's eye view moderate agreement and study models substantial agreement. On the other hand, the intra-class coefficient showed that both the frontal and worm's eye views had fair agreement and study models, excellent agreement.

Table IV – Inter-examiner agreement for frontal view, worm's eye view and study models.

	ICC (95% Confidence Interval)	Kappa
Frontal view	0.53 (0.42 to 0.63)	0.29
Worm's eye view	0.46 (0.33 to 0.57)	0.23
Study models	0.84 (0.79 to 0.88)	0.63

Table V - Inter-examiner agreement for frontal view, worm's eye view, and study models.

View	ICC (95% Confidence Interval)
Frontal view	0.52 (0.44 to 0.59)
Worm's eye view	0.52 (0.45 to 0.60)
Study models	0.86 (0.83 to 0.89)

Inter-rater reliability between all four raters was assessed using intra-class correlation coefficient (Table V). This showed that the agreement between study models was excellent. On the other hand, the frontal view and worm's eye view had fair agreement. Interestingly the level of agreement value was the same for both frontal and worm's eye view. On separating the results for consultants and senior registrars, worm's eye view was found to have a stronger agreement between consultants (Table III), and frontal views had a stronger agreement between senior registrars (Table IV).

DISCUSSION

This study was designed to assess if a relationship exists between the scores given to study models using the 5-Year Olds' Index (Atack et al., 1997b, Atack et al., 1997a), and scores given to the nasolabial appearance using frontal and worm's eye view photographs for the

same subjects on a 5 point Likert scale. In order to reduce memory bias, there was a two-week interlude before the photographs and study models were re-scored.

Intra-examiner agreement for frontal views, worm's eye views and study models showed that all 4 raters had excellent intra-class agreement when scoring study models. Scoring of the frontal and worm's eye view produced poorer levels of intra-class agreements. This may be associated with the fact that reference photographs were not available whilst scoring the frontal and worm's eye views. The inter- and intra-rater reliability of assessing nasolabial appearance has previously been reported as being poorer when compared to the reliability of assessing dental arch relationships using the Goslon Yardstick (Mercado *et al.*, 2016, Mars *et al.*, 1987). The Goslon Yardstick has been shown to be reliable and reproducible and has several example photos for each category, which are available during rating sessions (Mars *et al.*, 1987, Shaw *et al.*, 1992). Similarly the 5-Year Olds' Index (Atack *et al.*, 1997b, Atack *et al.*, 1997a) is also reproducible and reliable, but unlike Goslon reference models are available for each category as well as specific descriptors that the rater can use to assist in allocating a score to each case (Atack *et al.*, 1997b, Atack *et al.*, 1997a). Orthodontists are familiar with assessing models, and recently it has been shown that both digital photographs and 3D digital models are good substitutes to plaster models for scoring with the 5-Year Olds' Index (Chawla *et al.*, 2012). The important feature is therefore the presence of reference photographs, 3D images or plaster models during scoring. Similar "references" were not available in this study for raters assessing the frontal and worm's eye soft tissue images. A recent study by Mercado *et al.*, (2016) attempted to develop a yardstick of reference photographs for nasolabial appearance assessments that could be used on patients aged 5 to 7 years old with complete UCLP. The same authors Mercado *et al.*, (2011) had previously used the reference photographs (Kuijpers-Jagtman *et al.* (2009), but did not find a significant improvement in reliability scores. The scoring of the appearance of the nasolabial region and the 5-Year Olds' Index (Atack *et al.*, 1997b, Atack *et al.*, 1997a) is subjective. Arguably, the 5-Year Olds' Index (Atack *et al.*, 1997b, Atack *et al.*, 1997a) has slightly less subjectivity since reference criteria for each category are available.

Worm's eye view photographs scored worse than frontal views and this may be attributed to the fact that the examiners were not used to scoring this view. In addition, asymmetry is

also more noticeable in the worm's eye view than the frontal view (Gkantidis et al., 2013) and may have contributed to it being given worse scores. Profession and gender of the rater in addition to location of asymmetry can influence how asymmetries are perceived (Meyer-Marcotty et al., 2011, Johnston et al., 1999). Orthodontists have been found to be more critical than lay people, dental students and dental professionals when assessing chin asymmetry (McAvinchey et al., 2014, Naini et al., 2012). A recent review of the literature by Wang *et al.*, (2017) looked at asymmetries of the face and found that each facial feature has a unique threshold point at which asymmetry is perceived. The relationship in identifying asymmetry is not linear, as previously believed, but is of an exponential nature. The threshold point was identified when there was a sudden increase in detection rates of statistical significance. Eyelid position at rest is the most sensitive facial feature with a threshold of 2mm. With regards to the nasal tip, a deviation of more than 4mm is thought to be asymmetric by both lay people and clinicians (Wang et al., 2017).

The results show that the level of agreement between frontal and worm's eye view photographs for the same subjects was poor. Kim *et al.*, (2011) and Trotman (2013) looked at worm's eye view and frontal view photographs in the same studies, but neither compared the results for the two views for the same cleft subjects. Research by Al-Ghatam (2014) found poor agreement between frontal and worm's eye view photographs. Previous literature has shown that the appearance of the lip influences the combined lip and nose image, more than the appearance of the nose for the same image (Deall et al., 2016, Mosmuller et al., 2013). Shape and symmetry of the lips is easily assessed in the frontal view, whereas in the worm's eye view the lips do not appear as full. It is also known that some assessors score the lips first, and then adjust the score based on the condition of the nose (Deall et al., 2016). The difference in the presence of the lips in the two views may account for the discrepancy or lack of agreement between scores (Deall et al., 2016).

The raters used in this study were all orthodontists with varying levels of experience in the management of CL±P. Scoring of the study models with the 5-Year Olds' Index (Atack et al., 1997b, Atack et al., 1997a) meant that lay persons and other professionals could not be used, as they are not trained to use the index. Four raters were used in this study, but

within the literature there is a large variation in the number of raters used in different studies. Asher-McDade *et al.*, (1991) showed that by increasing the number of raters from three to six, the inter-examiner reliability increased from 0.83 to 0.9, therefore improving reliability and minimising inter-examiner bias. It has also been shown that a low number of raters can lead to issues with reliability (Marcusson *et al.*, 2002, Foo *et al.*, 2013). As a result, the literature has seen a large increase in the number of raters with studies using up to 39 raters (Eichenberger *et al.*, 2014, Papamanou *et al.*, 2012). However, this large increase in numbers has now been questioned by a recent study by Bella *et al.*, (2016), which showed that using 29 raters resulted in a lower inter- and intra-rater reliability score compared to those studies which had just four to six raters (Mercado *et al.*, 2011, Mercado *et al.*, 2016, Asher-McDade *et al.*, 1992).

The web-based scoring system developed by the Birmingham Institute of Paediatric Plastic Surgery, was used. It has been used in previous studies, including tri-centre audits (Deall *et al.*, 2016, Al-Ghatam *et al.*, 2015, Bella *et al.*, 2016).

The advantages of this online scoring system are:

- The security of the system, as each rater can only gain access using a personalised token.
- Scoring can be carried out at a time convenient to the raters.
- The system is simple to use.

Despite some of the advantages of this online scoring system, in the present study the 5-Year Olds' Index (Atack *et al.*, 1997b, Atack *et al.*, 1997a) was still found to be the most reliable method of assessing cleft treatment outcomes in this age group. Inter- and intra-rater reliability was found to be poor when scoring treatment outcomes using the frontal and worm's eye view photographs when compared to the study models. As a result, the correlation between the photographic assessments and occlusal assessments of the same 5-year-old cleft children's results was poor.

CONCLUSIONS

The following conclusions were reached concerning possible relationships between the aesthetic scores given to the nasolabial region of repaired UCLP patients and the 5-Year Olds' Index (Atack *et al.*, 1997b, Atack *et al.*, 1997a) scores for the same subjects, namely:

- There was very poor agreement observed between the scores given to the different photographic views (frontal and worm's eye) and their corresponding study models in the case of all the raters.
- The level of inter- and intra-rater reliability was strongest when scoring the study models using the 5-Year Olds' Index (Atack et al., 1997b, Atack et al., 1997a), suggesting this is still the most reliable outcome measure for this age group.

References

Al-Ghatam, R. 2014. Cleft Care Uk (CCUK)- Analysis of Photographs. Doctor of Dental Sciences in Orthodontics, University of Bristol, England.

Al-Ghatam, R., Jones, T. E. M., Ireland, A. J., Attack, N. E., Chawla, O., Deacon, S., Albery, L., Cobb, A. R. M., Cadogan, J., Leary, S., Waylen, A., Wills, A. K., Richard, B., Bella, H., Ness, A. R., Sandy, J. R.. Structural outcomes in the Cleft Care UK study. Part 2: dento-facial outcomes. *Orthodontics & Craniofacial Research*. 2015; 18: 14-24.

Al-Omari, I., Millett, D. T., Ayoub, A., Bock, M., Ray, A., Dunaway, D., Crampin, L.. An appraisal of three methods of rating facial deformity in patients with repaired complete unilateral cleft lip and palate. *Cleft Palate-Craniofacial Journal*. 2003;40: 530-537.

Asher-Mcdade, C., Brattstrom, V., Dahl, E., McWilliam, J., Molsted, K., Plint, D. A., Prahlandersen, B., Semb, G., Shaw, W. C.. A 6-Center International Study Of Treatment Outcome In Patients With Clefts Of The Lip And Palate .4. Assessment Of Nasolabial Appearance. *Cleft Palate-Craniofacial Journal*. 1992; 29: 409-412.

Asher-Mcdade, C., Roberts, C., Shaw, W. C., Gallager, C.. Development Of A Method For Rating Nasolabial Appearance In Patients With Clefts Of The Lip And Palate. *Cleft Palate-Craniofacial Journal*. 1991; 28: 385-390.

Assuncao, A. G. A.. The VLS Classification For Secondary Deformities In The Unilateral Cleft-Lip. *British Journal of Plastic Surgery*. 1992; 45: 288-292.

Attack, N., Hathorn, I., Mars, M., Sandy, J.. Study models of 5 year old children as predictors of surgical outcome in unilateral cleft lip and palate. *European Journal of Orthodontics*. 1997A; 19: 165-170.

Attack, N. E., Hathorn, I. S., Semb, G., Dowell, T., Sandy, J. R.. A new index for assessing surgical outcome in unilateral cleft lip and palate subjects aged five: Reproducibility and validity. *Cleft Palate-Craniofacial Journal*. 1997B; 34: 242-246.

Bella, H., Kornmann, N. S. S., Hardwicke, J. T., Wallis, K. L., Wearn, C., Su, T. L., Richard, B. M.. Facial aesthetic outcome analysis in unilateral cleft lip and palate surgery using web-based extended panel assessment. *Journal of Plastic Reconstructive and Aesthetic Surgery*. 2016; 69: 1537-1543.

Bongaarts, C. A. M., Prah-Andersen, B., Bronkhorst, E. M., Spauwen, P. H. M., Mulder, J. W., Vaandrager, J. M., Kuijpers-Jagtman, A. M.. Effect of infant orthopedics on facial appearance of toddlers with complete unilateral cleft lip and palate (Dutchcleft). *Cleft Palate-Craniofacial Journal*. 2008; 45: 407-413.

Brattstrom, V., McWilliam, J., Larson, O., Semb, G.. Craniofacial Development In Children With Unilateral Clefts Of The Lip, Alveolus, And Palate Treated According To 3 Different Regimes - Assessment Of Nasolabial Appearance. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*. 1992; 26: 313-319.

Brattstrom, V., Molsted, K., Prahl-Andersen, B., Semb, G. & Shaw, W. C. The Eurocleft study: Intercenter study of treatment outcome in patients with complete cleft lip and palate. Part 2: Craniofacial form and nasolabial appearance. *Cleft Palate-Craniofacial Journal*. 2005; 42: 69-77.

Chawla, O., Deacon, S. A., Attack, N. E., Ireland, A. J., Sandy, J. R.. The 5-year-olds' Index: determining the optimal format for rating dental arch relationships in unilateral cleft lip and palate. *European Journal of Orthodontics*. 2012; 34: 768-772.

Cicchetti, D. V.. Guidelines, criteria and rules of thumb of evaluating normal and standardised assessment instruments in psychology. *Psychological Assessment*. 1994; 6: 284-290.

Deall, C. E., Kornmann, N. S. S., Bella, H., Wallis, K. L., Hardwicke, J. T., Su, T. L., Richard, B. M. Facial Aesthetic Outcomes of Cleft Surgery: Assessment of Discrete Lip and Nose Images Compared with Digital Symmetry Analysis. *Plastic and Reconstructive Surgery*. 2016; 138: 855-862.

Edler, R., Abd Rahim, R., Wertheim, R., Greenhill, R.. The use of facial anthropometrics in aesthetic assessment. *Cleft Palate-Craniofacial Journal*. 2010; 47: 48-57.

Eichenberger, M., Staudt, C. B., Pandis, N., Gnoinski, W., Eliades, T. Facial attractiveness of patients with unilateral cleft lip and palate and of controls assessed by laypersons and professionals. *European Journal of Orthodontics*. 2014; 36: 284-289.

Eliason, M. J., Hardin, M. A., Olin, W. H.. Factors that influence ratings of facial Appearance for children with cleft-lip and palate. *Cleft Palate-Craniofacial Journal*. 1991; 28: 190-193.

Foo, P., Sampson, W., Roberts, R., Jamieson, L., David, D.. Facial aesthetics and perceived need for further treatment among adults with repaired cleft as assessed by cleft team professionals and laypersons. *European Journal of Orthodontics*. 2013; 35: 341-346.

Fudalej, P., Katsaros, C., Bongaarts, C., Dudkiewicz, Z., Kuijpers-Jagtman, A. M.. Dental arch relationship in children with complete unilateral cleft lip and palate following one-stage and three-stage surgical protocols. *Clinical Oral Investigations*. 2011; 15: 503-510.

Fudalej, P., Katsaros, C., Hozyasz, K., Borstlap, W. A., Kuijpers-Jagtman, A. M.. Nasolabial symmetry and aesthetics in children with complete unilateral cleft lip and palate. *British Journal of Oral & Maxillofacial Surgery*. 2012; 50: 621-625.

Gkantidis, N., Papamanou, D. A., Christou, P., Topouzelis, N.. Aesthetic outcome of cleft lip and palate treatment. Perceptions of patients, families, and health professionals compared to the general public. *Journal of Cranio-Maxillofacial Surgery*. 2013; 41: E105-E110.

Goodacre, T., Swan, M.C.. Cleft Lip and Palate: Current Management. *Paediatrics and Child Dental Health*. 2008; 18: 283-292.

Goodacre, T. E. E., Hentges, F., Moss, T. L. H., Short, V., Murray, L.. Does repairing a cleft lip neonatally have any effect on the longer-term attractiveness of the repair? *Cleft Palate-Craniofacial Journal*. 2004; 41: 603-608.

Howells, D. J., Shaw, W. C.. The validity and reliability of ratings of dental and

facial attractiveness for epidemiologic use. *American Journal of Orthodontics and Dentofacial Orthopedics*. 1985; 88: 402-408.

Huddart, A. G., Bodenham, R. S. The evaluation of arch form and occlusion in unilateral cleft palate subjects. *The Cleft Palate Journal*. 1972; 9: 194-209.

Jensen, M. P., Karoly, P. & Braver, S.. The measurement of clinical pain intensity - a comparison of 6 methods. *Pain*. 1986; 27: 117-126.

Johnson, N., Sandy, J.. An aesthetic index for evaluation of cleft repair. *European Journal of Orthodontics*. 2003; 25: 243-249.

Johnston, C. D., Burden, D. J., Stevenson, M. R.. The influence of dental to facial midline discrepancies on dental attractiveness ratings. *European Journal of Orthodontics*. 1999; 21: 517-522.

Jones, T., Al-Ghatam, R., Atack, N., Deacon, S., Power, R., Albery, L., Ireland, T., Sandy, J.. A review of outcome measures used in cleft care. *Journal of Orthodontics*. 2014; 41: 128-40.

Kalaaji, A., Lilja, J., Friede, H., Elander, A.. Bone grafting in the mixed and permanent dentition in cleft lip and palate patients: Long-term results and the role of the surgeon's experience. *Journal of Cranio-Maxillofacial Surgery*. 1996; 24: 29-35.

Kim, J. B., Strike, P., Cadier, M. C.. A simple assessment method for auditing multi-centre unilateral cleft lip repairs. *Journal of Plastic Reconstructive and Aesthetic Surgery*. 2011; 64: 195-200.

Knutsson, I., Rydstrom, H., Reimer, J., Nyberg, P., Hagell, P.. Interpretation of response categories in patient-reported rating scales: a controlled study among people with Parkinson's disease. *Health and Quality of Life Outcomes*. 2010; 8: 61

Kuijpers-Jagtman, A. M., Nollet, P. J. P. M., Semb, G., Bronkhorst, E. M., Shaw, W. C., Katsaros, C.. Reference Photographs for Nasolabial Appearance Rating in Unilateral Cleft Lip and Palate. *Journal of Craniofacial Surgery*. 2009; 20: 1683-1686.

Kwak, K. H., Kim, Y. I., Nam, H. J., Kim, S. S., Park, S. B., Son, W. S. Differences among deviations, genders, and observers in the perception of eye and nose asymmetry. *Journal of Oral and Maxillofacial Surgery*. 2015; 73: 1606-1614.

Landis, J. R., Koch, G. G.. Measurement of observer agreement for categorical Data. *Biometrics*. 1977; 33: 159-174.

Mani, M. R., Semb, G., Andlin-Sobocki, A.. Nasolabial appearance in adults with repaired unilateral cleft lip and palate: Relation between professional and lay rating and patients' satisfaction. *Journal of Plastic Surgery and Hand Surgery*. 2010; 44: 191-198.

Marcusson, A., Paulin, G., Ostrup, L.. Facial appearance in adults who had cleft lip and palate treated in childhood. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*. 2002; 36: 16-23.

Mars, M., Plint, D. A., Houston, W. J. B., Bergland, O., Semb, G.. The Goslon Yardstick - a new system of assessing dental arch relationships in children with Unilateral clefts of the lip and palate. *Cleft Palate Journal*. 1987; 24: 314-322.

Mcavinchey, G., Maxim, F., Nix, B., Djordjevic, J., Linklater, R., Landini, G.. The perception of facial asymmetry using 3-dimensional simulated images. *Angle Orthodontist*. 2014; 84: 957- 965.

Mercado, A., Russell, K., Hathaway, R., Daskalogiannakis, J., Sadek, H., Long, R. E., Cohen, M., Semb, G., Shaw, W.. The Americleft Study: An Inter-Center Study of Treatment Outcomes for Patients With Unilateral Cleft Lip and Palate Part 4. Nasolabial Aesthetics. *Cleft Palate-Craniofacial Journal*. 2011; 48: 259-264.

Mercado, A. M., Russell, K. A., Daskalogiannakis, J., Hathaway, R. R., Semb, G., Ozawa, T., Smith, A., Lin, A. Y., Long, R. E.. The Americleft Project: A Proposed Expanded Nasolabial Appearance Yardstick for 5-to 7-Year-Old Patients With Complete Unilateral Cleft Lip and Palate (CUCLP). *Cleft Palate-Craniofacial Journal*. 2016; 53: 30-37.

Meyer-Marcotty, P., Gerdes, A. B. M., Reuther, T., Stellzig-Eisenhauer, A., Alpers, G. W.. Persons with Cleft Lip and Palate Are Looked at Differently. *Journal of Dental Research*. 2010; 89: 400-404.

Meyer-Marcotty, P., Stellzig-Eisenhauer, A.. Dentofacial Self-Perception and Social Perception of Adults with Unilateral Cleft Lip and Palate. *Journal of Orofacial Orthopedics-Fortschritte Der Kieferorthopadie*. 2009; 70: 224-236.

Meyer-Marcotty, P., Stellzig-Eisenhauer, A., Bareis, U., Hartmann, J., Kochel, J.. Three-dimensional perception of facial asymmetry. *European Journal of Orthodontics*. 2011; 33: 647-653.

Morrant, D. G., Shaw, W. C.. Use of standardized video recordings to assess cleft surgery outcome. *Cleft Palate-Craniofacial Journal*. 1996; 33: 134-142.

Mosmuller, D. G. M., Bijnen, C. L., Kramer, G. J. C., Disse, M. A., Prahl, C., Kuik, D. J., Niessen, F. B., Griot, J.. The Asher-McDade Aesthetic Index in Comparison With Two Scoring Systems in Nonsyndromic Complete Unilateral Cleft Lip and Palate Patients. *Journal of Craniofacial Surgery*. 2015; 26: 1242-1245.

Mosmuller, D. G. M., Griot, J. P. W. D., Bijnen, C. L., Niessen, F. B.. Scoring Systems of Cleft-Related Facial Deformities: A Review of Literature. *Cleft Palate-Craniofacial Journal*. 2013; 50: 286-296.

Mossey, P. A., Clark, J. D., Gray, D.. Preliminary investigation of a modified Huddart/Bodenham scoring system for assessment of maxillary arch constriction in unilateral cleft lip and palate subjects. *European Journal of Orthodontics*. 2003; 25: 251-257.

Mossey, P. A., Little, J., Munger, R. G., Dixon, M. J. & Shaw, W. C.. Cleft lip and palate. *Lancet*. 2009; 374: 1773-1785.

Naini, F. B., Donaldson, A. N. A., Cobourne, M. T., McDonald, F.. Assessing the influence of mandibular prominence on perceived attractiveness in the orthognathic patient, clinician, and layperson. *European Journal of Orthodontics*. 2012; 34: 738-746.

Ogon, M., Krismer, M., Sollner, W., Kantnerrumplmaier, W., Lampe, A.. Chronic low

back pain measurement with visual analogue scales in different settings. *Pain*. 1996; 64: 425-428.

Ohannessian, P., Berggren, A., Abdiu, A.. The cleft lip evaluation profile (CLEP): A new approach for postoperative nasolabial assessment in patients with unilateral cleft lip and palate. *Journal of Plastic Surgery and Hand Surgery*. 2011; 45: 8-13.

Okkerse, J. M. E., Beemer, F. A., Cordia-De Boer, M., Heineman-De Boer, J. A., Mellenbergh, G. J., Wolters, W. H. G.. Facial attractiveness and facial impairment ratings in children with craniofacial malformations. *Cleft Palate-Craniofacial Journal*. 2001; 38: 386-392.

Papamanou, D. A., Gkantidis, N., Topouzelis, N., Christou, P.. Appreciation of cleft lip and palate treatment outcome by professionals and laypeople. *European Journal of Orthodontics*. 2012; 34: 553-560.

Persson, M., Sandy, J. R., Waylen, A., Wills, A. K., Al-Ghatam, R., Ireland, A. J., Hall, A. J., Hollingworth, W., Jones, T., Peters, T. J., Preston, R., Sell, D., Smallridge, J., Worthington, H., Ness, A. R.. A cross-sectional survey of 5-year-old children with non-syndromic unilateral cleft lip and palate: the Cleft Care UK study. Part 1: background and methodology. *Orthodontics & Craniofacial Research*. 2015; 18: 1-13.

Pigott, R. W., Pigott, B. B.. Quantitative Measurement of Symmetry From Photographs Following Surgery for Unilateral Cleft Lip and Palate. *Cleft Palate-Craniofacial Journal*. 2010; 47: 363-367.

Prahl, C., Prahl-Andersen, B., Van 'T Hof, M. A., Kuijpers-Jagtman, A. M.. Infant orthopedics and facial appearance: A randomized clinical trial (Dutchcleft). *Cleft Palate-Craniofacial Journal*. 2006; 43: 659-664.

Richmond, S., Turbill, E. A., Andrews, M.. Calibration of non-dental and dental personnel in the use of the PAR Index. *British journal of orthodontics*. 1993; 20: 231-4.

Roberts, C. T., Semb, G., Shaw, W. C.. Strategies for the advancement of surgical Methods in cleft-lip and palate. *Cleft Palate-Craniofacial Journal*. 1991; 28: 141-149.

Sandy, J., Rumsey, N., Persson, M., Waylen, A., Kilpatrick, N., Ireland, T., Ness, A.. Using service rationalisation to build a research network: lessons from the centralisation of UK services for children with cleft lip and palate. *British Dental Journal*. 2012; 212: 553-555.

Scott, J. K., Leary, S. D., Ness, A. R., Sandy, J. R., Persson, M., Kilpatrick, N., Waylen, A. E.. Centralization of Services for Children Born With Orofacial Clefts in the United Kingdom: A Cross-Sectional Survey. *Cleft Palate-Craniofacial Journal*. 2014; 51: E102-E109.

Shaw, W. C.. The influence of childrens dentofacial appearance on their social Attractiveness as judged by peers and lay adults. *American Journal of Orthodontics and Dentofacial Orthopedics*. 1981; 79: 399-415.

Shaw, W. C., Asher-Mcdade, C., Brattstrom, V., Dahl, E., Mcwilliam, J., Molsted, K., Plint, D. A., Prahlandersen, B., Semb, G., The, P. S. R.. A 6-center international study of Treatment outcome in patients with clefts of the lip and palate .1. Principles and Study design. *Cleft Palate-Craniofacial Journal*. 1992; 29: 393-397.

Shaw, W. C., Rees, G., Dawe, M., Charles, C. R.. The influence of dentofacial appearance on the social attractiveness of young-adults. *American Journal of Orthodontics and Dentofacial Orthopedics*. 1985; 87: 21-26.

Snow, S., Kirwan, J. R.. Visual analog scales - a source of error. *Annals of the Rheumatic Diseases*. 1988; 47: 526-526.

Stebel, A., Desmedt, D., Bronkhorst, E., Kuijpers, M. A., Fudalej, P. S.. Rating nasolabial appearance on three-dimensional images in cleft lip and palate: a comparison with standard photographs. *European Journal of Orthodontics*. 2016; 38: 197-201.

Svensson, E.. Concordance between ratings using different scales for the same variable. *Statistics in Medicine*. 2000; 19: 3483-3496.

Tedesco, L. A., Albino, J. E., Cunat, J. J., Green, L. J., Lewis, E. A., Slakter, M. J. A.. Dental-facial attractiveness scale .1. Reliability and validity. *American Journal of Orthodontics and Dentofacial Orthopedics*. 1983; 83: 38-43.

Tobiasen, J. M.. Social judgments of facial deformity. *Cleft Palate Journal*. 1987; 24: 323-327.

Trotman, C. A., Phillips, C., Faraway, J. J., Hartman, T., Van Aalst, J. A.. Influence of Objective Three-Dimensional Measures and Movement Images on Surgeon Treatment Planning for Lip Revision Surgery. *Cleft Palate-Craniofacial Journal*. 2013; 50: 684-695.

Vegter, F., Hage, J. J.. Standardized facial photography of cleft patients: Just fit the grid? *Cleft Palate-Craniofacial Journal*. 2000; 37: 435-440.

Wang, T. T., Wessels, L., Hussain, G., Merten, S.. Discriminative Thresholds in Facial Asymmetry: A Review of the Literature. *Aesthetic Surgery Journal*. 2017; 37: 375-385.

Williams, A. C., Bearn, D., Mildinhall, S., Murphy, T., Sell, D., Shaw, W. C., Murray, J. J., Sandy, J. R.. Cleft lip and palate care in the United Kingdom - The Clinical Standards Advisory Group (CSAG) study. Part 2: Dentofacial outcomes and patient satisfaction. *Cleft Palate-Craniofacial Journal*. 2001; 38: 24-29.

Williams, A. C., Sandy, J. R., Thomas, S., Sell, D., Sterne, J. A. C.. Influence of surgeon's experience on speech outcome in cleft lip and palate. *Lancet*. 1999; 354: 1697-1698.

Witt, P. D., Wahlen, J. C., Marsh, J. L., Grames, L. M., Pilgram, T. K.. The effect of surgeon experience on velopharyngeal functional outcome following palatoplasty: Is there a learning curve? *Plastic and Reconstructive Surgery*. 1998; 102: 1375-1384.